Journal of Forestry Research (2009) 20(3): 195−198
DOI 10.1007/s11676-009-0038-7

RESEARCH PAPER

Development of two microsatellite multiplex PCR systems for high throughput genotyping in *Populus euphratica*

Eusemann Pascal*, Fehrenz Steffen, Schnittler Martin
Department of Botany and Landscape Ecology, EMAU Greifswald, Grimmer Str. 88, 17487 Greifswald, Germany

Abstract: Eighteen microsatellite primer pairs previously developed at Oak Ridge National Laboratory for *Populus tremuloides* Michx. and *Populus trichocarpa* Torr. & Gray were screened for amplification in Euphrates poplar, *Populus euphratica* Oliv. Thirteen loci were found to express polymorphisms ranging from two to 17 alleles. The eight most variable loci were selected to set up and optimize two multiplex polymerase chain reaction (PCR) assays. Three populations containing altogether 436 trees were used to characterize the selected loci and ascertain their applicability for parentage analysis and genotyping studies. Through cross-checking of clonal identity against sex of the genotyped trees we estimated the maximum error rate for merging genotypes to be less than 0.045.

Keywords: clone identification; Euphrates poplar; genetic fingerprint; parentage analysis; population structure; SSR primers

Introduction

Euphrates poplar (*Populus euphratica* Oliv.) is an important tree species of riparian ecosystems in arid Central Asia (Wang et al. 1996). In the river corridors of north-western China it plays important ecological functions as the keystone species for all biodiversity in the region (Thevs 2005). Being the main source of wood and an important crop plant, *P. euphratica* also has considerable economic value for the local human population and suffers from strong pressure by humans and their livestock (Wang et al. 1996; Weisgerber et al. 1995). Within the genus, *P. euphratica* and the closely related species *P. pruinosa* Schrenk form an own section, Turanga. The dioecious tree is able to reproduce sexually and by clonal growth via root suckering.

Despite the species' ecological, economic, and conservational importance, only few studies addressed population structure, population dynamics, or demographic processes in natural stands and those typically employed anonymous markers such as AFLP and RAPD (Bruelheide et al. 2004; Fay et al. 1999; Saito et al. 2002). Microsatellite markers developed specifically for this species were recently reported by Wu et al. (2008). To apply this powerful marker system (see Selkoe & Toonen 2006 for review) on Euphrates poplar, we had simultaneously tested primers developed at Oak Ridge National Laboratory for black cottonwood (*P. trichocarpa* Torr. & Gray) and aspen (*P. tremuloides* Michx.) for amplification in *P. euphratica* and characterised the successfully transferred loci on the basis of three natural populations containing trees produced both sexually and clonally.

In this paper, we present the results of primer screening, the characteristics of eight loci suitable to address population genetic questions, and the subsequent development of two multiplex PCR systems. The use of multiplex PCR increases work flow efficiency while simultaneously reducing per sample cost (Vaughan & Russel 2004), thus considerably facilitating large scale population studies. The aim of this study is to provide researchers with an accurate and robust tool to assess the genetic diversity of large sample numbers of this important forest tree.

Material and Methods

Primer screening

Primers were chosen from the literature (Tuskan et al. 2004) and from internet resources of Oak Ridge National Laboratory. Of the almost 4 200 primer pairs listed such primers were selected that showed a high number of alleles in natural aspen populations, and/or amplified PCR-products in different size categories to facilitate the following development of multiplex PCR assays for these primers. In total, 18 primer pairs were used to pre-screen 220 trees from different stands at the middle reaches of the Tarim River, Xinjiang, China (N41°12'13", E84°22'56") and 16 plants from Azerbaijan (N40°50'40", E49°17'45"). Each primer pair was tested in a separate PCR.

Plants selected

All selected loci were subsequently characterised in 145, 158 and 133 trees of three populations called Ing5, Ing6 and Ing8, located

Foundation project: This study was supported by the Deutsche Forschungsgemeinschaft DFG (grant number SCHN 1080/1-1)
Received: 2009-03-06; Accepted: 2009-04-05

The online version is available at http://www.springerlink.com
Biography: Pascal Eusemann, Corresponding author, born 1981.02.18, male, PhD student at EMAU Greifswald, research interests are in forest ecology and forest genetics. Phone: 0049-(0)-3834864111. Fax: 0049-(0)-3834864114. Email: pascal.eusemann@uni-greifswald.de
Senior author: Martin Schnittler. Phone: 0049-(0)-3834864123. Fax: 0049-(0)-3834864114. Email: martin.schnittler@uni-greifswald.de
Responsible editor: Chai Ruihai

at the middle reaches of the Tarim River (Ing6: N41°13'51", E84°12'18"). The distance between stands Ing5 and Ing6 is 400 m and between Ing8 and the Ing5/Ing6 complex about 2,500 m. All stands were mapped with a differential GPS (Trimble R3, precision in floating mode 0.1 m) during flowering time to determine the sex of the flowering trees. Stand characteristics are shown in Table 1. About 50 per cent of all sampled trees were selected for genetic analysis using the grid sampling method proposed by Suzuki et al. (2004) for recovering a maximum of genotypes with a minimum of samples. Basically, each tree located closest to a mesh of a rectangular grid was chosen to achieve an even distribution of analysed trees over the stand.

Table 1. Characteristics of the studied stands of *Populus euphratica*.

Stand	Trees sampled	Trees analysed	Number of genotypes	Number of clones
Ing5	339	145	56	20
Ing6	249	158	54	20
Ing8	260	133	90	16
Stand	Total trees in clones	Mean clone size [trees]	Proportion of clones	
Ing5	109	5.45	0.75	
Ing6	124	6.2	0.78	
Ing8	59	3.69	0.44	

Population genetic analysis

Values for Hardy-Weinberg equilibrium (HWE), linkage disequilibrium (LD) (GENEPOP 4.0; Rousset 2008), observed and expected heterozygosities, null allele frequencies, probability of identity (P_{ID}) (IDENTITY 1.0; Wagner & Sefc 1999), and exclusion probabilities for parentage analysis (FaMoz; Gerber et al. 2003) were calculated for every locus on the basis of total trees and genets (only one tree per genotype included in the calculation). P_{ID} and exclusion probabilities were furthermore calculated over all loci. Calculations were carried out for each population independently and over all populations.

The resolution of the genotyping method was verified by comparing the sex of trees within a clone. Sex could be ascertained for 300 of the 436 trees studied. Trees which differed in sex from the remaining trees of the clone they were assigned to were counted as errors. For a dioecious species with even sex ratio, a sex error has a probability of 50% to be detected if the sex of a tree is known. Therefore, the real number of errors is twice that high and depends in addition from the proportion of trees with known sex. Accounting for the skewed sex ratios as well, the maximum error E for merging genotypes is thus

$$E = E_{(sex)} * 1/r * N_{(analysed\ trees)} / N_{(sex\text{-}determined\ trees)}$$

where, $E_{(sex)}$ is the number of sex errors (trees with a sex deviating from the rest of the clone), r is proportion of the rarer sex, $N_{(analysed\ trees)}$ is the number of trees included in the study, and $N_{(sex\text{-}determined\ trees)}$ is the number of trees with known sex in the study.

Multiplex PCR

Two multiplex PCR, each amplifying four loci, were developed. All primers were tested for possible primer-primer interactions and hairpin structures using AutoDimer software 1.0 (Vallone & Butler 2004). Primer combinations and characteristics are given in Table 1. A total PCR volume of 6 μL, containing 3.5 ng DNA, 1x PCR buffer (Qiagen), 0.2 mM dNTP Mix (Fermentas), 0.3 U *Taq* polymerase (Molzym), 0.4 mM BSA (New England Biolabs), and primer concentrations according to Table 1 was used. Final primer concentrations were adjusted empirically to homogenize loci amplification, starting with equimolar concentrations of 0.4 μM and successively reducing primer concentration of stronger loci to obtain balanced signals. PCRs were performed on Eppendorf Mastercycler thermocyclers under the following conditions: A cooling step of 5 min duration at 4°C while the lid of the thermocycler heats up, a denaturation step of 5 min at 94°C, followed by 30 cycles of 94°C for 30 s, annealing at 60°C, 58°C, 56°C, 54°C, 52°C, and 50°C for 10 s each, 72°C for 30 s, and a singular extension of 45 min at 60°C. Subsequently, the samples were cooled at 21°C for 10 min before cooling down to a final temperature of 4°C. This profile meets the specific annealing temperature for every primer in the reaction mixture. Though other cycle conditions (like a classic touch down profile or the use of only one annealing temperature) were tested during optimization as well, our procedure was found to be most effective in reducing unspecific binding and background noise. The final extension step is recommended to reduce stutter phenomena and increase PCR efficiency in multiplex PCR (Henegariu et al. 1997; Lepais et al. 2006). PCR products were diluted 1:20 in water. For analysis, 1 μL of diluted product was combined with 0.1 μL GeneScan 500 Rox size standard (Applied Biosystems) and 8.9 μL HiDi Formamide (Applied Biosystems). Fragment analysis was carried out on an ABI Prism 310 capillary sequencer (Applied Biosystems). Genotyping was performed using GeneMapper 3.7 (Applied Biosystems).

Results

Primer screening

Of the 18 primer pairs studied, 13 generated polymorphic amplification products (72.2%), four were found to be monomorphic and one did not show any amplification at all. Between two and 17 alleles (mean 6.4) were revealed for the 13 polymorphic loci. The eight most polymorphic loci, displaying between four and 17 alleles (mean 9.1), were selected for further use in the genotyping approach. The ten loci discarded from further characterization were GCPM 2341 (no amplification), GCPM 3077 (1 allele), GCPM 3333 (1), ORPM 011 (2), ORPM 014 (2), ORPM 020 (2), ORPM 021 (1), ORPM 026 (2), ORPM 055 (2), and ORPM 1422 (1).

Population genetic analysis

Primer characteristics are shown in Table 2. Results of tests for expected and observed heterozygosity and exclusion probabilities for parentage analysis on the genet level for all three populations are shown in Table 3. Results of tests for Hardy-Weinberg equilibrium and pairwise tests for linkage disequilibrium for each

single population revealed significant deviation from HWE for five to seven loci. We found significant LD for all pairs of loci in Ing5 and Ing6, and for 23 out of 28 pairs of loci in Ing8 if calculated on the basis of all trees (p<0.05 for both tests). Calculation on the basis of genets showed significant deviation from HWE for five loci in all three populations. Significant LD was found for 11–27 pairs of loci (p<0.05). P_{ID} computed over all populations was 2.39 x 10^{-5} for all trees and 1.81 x 10^{-5} for all genets. P_{ID} ranged from 4.49×10^{-5}–2.47×10^{-4} in each single population for all trees and from 4.15×10^{-5}–1.22×10^{-4} for all genets. Estimated null allele frequencies were close to 0 for all loci in each population except for locus ORPM 016, for which frequencies ranged from 0.12–0.32 in the three populations studied. Cumulated exclusion probabilities over all eight loci and all three populations were 0.89 (single parent), 0.98 (paternity), and 1.00 (parent pair) both for all trees sampled and all genets. Values after exclusion of ORPM 016 were 0.87 (single parent), 0.97 (paternity), and 1.00 (parent pair).

Table 2. Characteristics of the eight microsatellite loci used in two multiplex PCRs

	Locus	Primer Sequence	Conc. (μM)	Dye	Allele Number	Size Range (bp)
Set 1	ORPM 016	F: 5'-GCAGAAACCACTGCTAGATGC-3' R: 5'-GCTTTGAGGAGGTGTGAGGA-3'	0.25	Tamra	4	217–226
	ORPM 1249	F: 5'-ACCTAAGGGTTGGAAGGTAG-3' R: 5'-CCCAAATGAAAAACAAAAGA-3'	0.20	Tamra	10	103–121
	ORPM 1261	F: 5'-TGCAGTTCTCCATGAACATA-3' R: 5'-GAAGTTTTTGACCTGCAGAC-3'	0.05	Hex	4	123–129
	GCPM 3351	F: 5'-AACCTCCAATACCAAGATCA-3' R: 5'-TGAGAATAAATATTTCGGCAA-3'	0.40	Fam	17	174–206
Set 2	ORPM 023	F: 5'-ATTCCATTTGGCAATCAAGG-3' R: 5'-CCCTGAAAGTCACGTCTTCG-3'	0.16	Tamra	9	197–215
	ORPM 030	F: 5'-ATGTCCACACCCAGATGACA-3' R: 5'-CCGGCTTCATTAAGAGTTGG-3'	0.04	Hex	12	207–229
	ORPM 1031	F: 5'-ATGTTTCGTATTTGGAATGG-3' R: 5'-GGCTTGGACTAGAGATGATG-3'	0.04	Fam	6	104–122
	GCPM 2768	F: 5'-TTATTTGGATCCTGAAATGG-3' R: 5'-GATGGTTCGGTATGTGAGTT-3'	0.11	Hex	11	175–195

Table 3. Population genetic characteristics of the eight loci calculated for each of the three populations examined

Locus	Stand Ing5					Ing6					Ing8				
	Exclusion probability			Heterozygosity		Exclusion probability			Heterozygosity		Exclusion probability			Heterozygosity	
	sp	p	pp	He	Ho	sp	p	pp	He	Ho	sp	p	pp	He	Ho
ORPM 016	0.837	0.968	0.997	0.713	0.132	0.831	0.962	0.996	0.528	0.342	0.543	0.706	0.877	0.711	0.280
ORPM 1249	0.819	0.960	0.996	0.401	0.453	0.844	0.971	0.998	0.238	0.240	0.653	0.822	0.945	0.357	0.376
ORPM 1261	0.842	0.972	0.998	0.253	0.208	0.839	0.967	0.997	0.302	0.280	0.724	0.885	0.973	0.321	0.344
GCPM 3351	0.439	0.616	0.800	0.803	0.736	0.450	0.626	0.811	0.807	0.820	0.769	0.919	0.985	0.843	0.882
ORPM 023	0.600	0.792	0.927	0.692	0.547	0.805	0.947	0.993	0.581	0.458	0.798	0.935	0.989	0.575	0.355
ORPM 030	0.754	0.919	0.985	0.627	0.623	0.613	0.804	0.938	0.687	0.620	0.819	0.950	0.993	0.523	0.505
ORPM 1031	0.688	0.868	0.965	0.629	0.585	0.761	0.922	0.986	0.595	0.600	0.832	0.961	0.996	0.683	0.634
GCPM 2768	0.791	0.943	0.992	0.653	0.698	0.704	0.880	0.972	0.675	0.700	0.838	0.966	0.997	0.535	0.301

Note: sp, single parent; p, paternity; pp, parent pair; H_e, expected heterozygosity; H_o, observed heterozygosity.

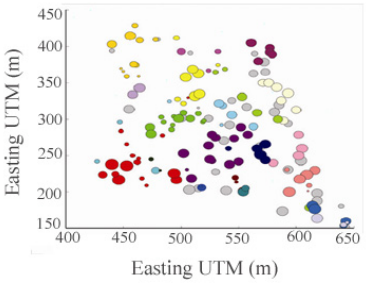

Fig. 1 Diameter at breast height and clonal identity of 158 trees analysed for the population Ing6 at the middle reaches of the Tarim River. Grey circles depict single genotypes. The twenty clones found in this stand are indicated by different colours.

Added over all populations, 56 genotypes were clones containing a total of 292 trees. Fig.1 shows the spatial distribution of genotypes for stand Ing6. Characteristics of all stands are shown in Table 1. In six cases (four in Ing5, two in Ing6) one tree in a clone had a sex differing from the rest of the trees in the clone. The estimated maximum error for merging clones was calculated as 13.5, 6.1 and 0.0 for the populations Ing5, Ing6 and Ing8, respectively. This transfers to a maximum error rate per analysed tree of 0.045 for all three populations together.

Discussion

For establishing a molecular method to genotype individuals, the resolution (in terms of the ability of the method to distinguish between genotypes) is of crucial importance. The P_{ID} values of the eight loci presented in this paper, derived both from every single population and from all populations combined, are deemed sufficient to distinguish even siblings with high confidence (Hoffman & Amos 2005). Null alleles are considered to have a negligible impact on genotyping experiments but may severely compromise parentage analysis due to false exclusion of true parents (Dakin & Avise 2004). Exclusion of locus ORPM 016 from the primer set led to P_{ID} values that still allow even siblings to be distinguished with high confidence. Exclusion probabilities for parentage analysis also remained high. While the locus is used in genotyping studies, it is possible to exclude it from the set when attempting parentage analysis with the remaining seven loci still being able to reliably perform genotyping and

parentage assignment.

The analysis of sex errors (trees with opposite sex in a clone) also indicates a sufficiently high resolution of the two microsatellite sets. Sex errors can be caused by genotyping errors but also by a wrong determination of a tree's sex in the field. Sexing old and only sparsely flowering trees in the field can be difficult. This was the case in the stands Ing5 and Ing6, containing old (usually >80 yrs) trees with low vitality. Hence, the maximum error rate of 0.045 obtained is a conservative calculation but still among the error rates reported by Hoffman & Amos (2005). A low error rate is also suggested by the spatial distribution of trees assigned to one genotype (Fig. 1).

Most of our loci deviated from Hardy-Weinberg proportions and showed linkage disequilibrium. Wu et al. (2008) presented 12 new microsatellite loci for *Populus euphratica* as well and report these to be in Hardy-Weinberg equilibrium and in linkage equilibrium. The apparent contradiction to our results is most likely explained by an error in interpreting test results. The authors report *p*-values smaller than 0.01 and 0.05 for the tests for HWE and LD, respectively. However, null hypotheses for the tests implemented in GENEPOP on the web (version 3.4, used by these authors) are that loci are in HWE and linkage equilibrium. *p*-values below 0.05 thus indicate significant deviation from HWE and significant LD. Reading the figures of Wu et al. (2008) this way, their results are in accordance with those obtained by us.

Deviations from HWE and LD can result from different causes both artificial and natural. In organisms that employ both sexual and asexual reproduction such deviations are characteristic life history features (Halkett et al. 2005). Our findings that LD and deviations from HWE are still maintained even if all replicate genotypes were eliminated from the data set is in accordance with the results of a study on *Prunus avium*, an also partially clonal tree species (Stoeckel et al. 2006).

The set of eight microsatellite loci for genotyping and parentage analysis in *P. euphratica* presented here was demonstrated to reliably identify individual genotypes in natural populations with a low error rate. It was proven to be sufficient for fine scale population genetic studies and parentage assignment as well. The two PCR multiplex sets are especially well suited for large scale population studies. An ongoing project studying the genetic structure of Euphrates poplar forests in NW China aims for the analysis of about 1 500 plants. For these numbers, using the two multiplex kits presented in this paper will save about 9000 PCR reactions and hence considerably reduce per sample costs and processing time.

Acknowledgements

We thank Yin TongMing, Oak Ridge National Laboratory, Oak Ridge, TN, USA for providing primer information and help in primer selection. This study was conducted in cooperation with the Institute of Natural Resources and Environmental Sciences of the Xinjiang University (group Prof. Nurbay Abdusalih). The project was supported by the Deutsche Forschungsgemeinschaft (DFG, grant number SCHN 1080/1-1).

References

Bruelheide H, Manegold M, Jandt U. 2004. The genetical structure of *Populus euphratica* and *Alhagi sparsifolia* stands in the Taklimakan desert. In: M. Runge and X. Zhang (eds), *Ecophysiology and habitat requirements of perennial plant species in the Taklimakan desert.* Contributions to a "Workshop on sustainable management of the shelterbelt vegetation of river oases in the Taklimakan desert". Urumqi, pp. 153–160.

Dakin EE, Avise JC. 2004. Microsatellite null alleles in parentage analysis. *Heredity*, **93**(5): 504–509.

Fay MF, Lledó MD, Kornblum MM, Crespo MB. 1999. From the waters of Babylon? *Populus euphratica* in Spain is clonal and probably introduced. *Biodiversity and Conservation*, **8**(6): 769–778.

Gerber S, Chabrier P, Kremer A. 2003. FaMoz: a software for parentage analysis using dominant, codominant and uniparentally inherited markers. *Molecular Ecology Notes*, **3**(4): 479–481.

Halkett F, Simon J, Balloux F. 2005. Tackling the population genetics of clonal and partially clonal organisms. *Trends in Ecology and Evolution*, **20**(4): 194–201.

Henegariu O, Heerema NA, Dlouhy SR, Vance GH, Vogt PH. 1997. Multiplex PCR: critical parameters and step-by-step protocol. *BioTechniques*, **23**(3): 504–511.

Hoffman JI, Amos W. 2005. Microsatellite genotyping errors: detection approaches, common sources and consequences for paternal exclusion. *Molecular Ecology*, **14**(2): 599–612.

Lepais O, Léger V, Gerber S. 2006. Short Note: high throughput microsatellite genotyping in oak species. *Silvae Genetica*, **55**(4/5): 238–240.

Rousset F. 2008. GENEPOP'007: a complete re-implementation of the GENEPOP software for Windows and Linux. *Molecular Ecology Resources*, **8**(1): 103–106.

Saito Y, Shiraishi S, Tanimoto T, Yin L, Watanabe S, Ide Y. 2002. Genetic diversity of *Populus euphratica* populations in northwestern China determined by RAPD DNA analysis. *New Forests*, **23**(2): 97–103.

Selkoe KA, Toonen RJ. 2006. Microsatellites for ecologists: a practical guide to using and evaluating microsatellite markers. *Ecology Letters*, **9**(5): 615–629.

Stoeckel S, Grange J, Fernández-Manjarres JF, Bilger I, Frascaria-Lacoste N, Mariette S. 2006. Heterozygote excess in a self-incompatible and partially clonal tree species, *Prunus avium* L. *Molecular Ecology*, **15**(8): 2109–2118.

Suzuki J, Herben T, Maki M. 2004. An under-appreciated difficulty: sampling of plant populations for analysis using molecular markers. Evolutionary Ecology, **18**(5–6): 625–646.

Thevs N. 2005. Tugay vegetation in the middle reaches of the Tarim River. Vegetation types and their ecology. *Archiv für Naturschutz und Landschaftsforschung*, **44**(1): 63–84.

Tuskan GA, Gunter LE, Yang ZK, Yin T, Sewell MM, DiFazio SP. 2004. Characterization of microsatellites revealed by genomic sequencing of *Populus trichocarpa*. *Canadian Journal of Forest Research*, **34**(1): 85–93.

Vallone PM, Butler JM. 2004. AutoDimer: a screening tool for primer-dimer and hairpin structures. *BioTechniques*, **37**(2): 226–231.

Vaughan SP, Russell K. 2004. Characterization of novel microsatellites and development of multiplex PCR for large-scale population studies in wild cherry, *Prunus avium*. *Molecular Ecology Notes*, **4**(3): 429–431.

Wagner HW, Sefc KM. 1999. IDENTITY 1.0. Centre for Applied Genetics, University of Agricultural Sciences Vienna.

Wang S, Chen B, Li H. 1996. Euphrates poplar forest. Beijing: China Environmental Science Press, 212 pp.

Weisgerber H, Kownatzki D, Mussong M. 1995. Natural poplar resources in China and their significance for breeding and afforestation. *Silvae Genetica*, **44**(4): 298–303.

Wu Y, Wang J, Liu J. 2008. Development and characterization of microsatellite markers in *Populus euphratica* (Populaceae). *Molecular Ecology Resources*, **8**(5): 1142–1144.